DOCUMENT RESUME

ED 197 710 IR 009 096

AUTHOR Beal, Christina
TITLE The Users of Indexes.
PUB DATE Sep 80
NOTE 24p.: Paper presented at the ASLIB/IIS/LA Joint Conference (Sheffield, England, September 15-19, 1980).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Indexes; *Indexing; Information Seeking; *Information Systems; Library Research; *User Satisfaction (Information); *Use Studies
IDENTIFIERS *Library Users

ABSTRACT

Problems of indexing and index use lie at the heart of information use. They determine maximum possible use of information systems and sources, affect communication with users, influence the use of retrieved documents, and should form a more important field of research than past history indicates. Previous research provides some clues to user behavior and user needs: however, indexing systems are sometimes uncritically accepted as unproblematic for the user. Some work has tackled the user problem by trying to structure index systems from the user's own formulations of inquiries, but there is a strong feeling that more instruction is needed in schools to teach children how to make full use of indexes and index systems, to make better use of index skills in everyday life, and to view indexes as diverse ways of structuring knowledge and information. This paper presents examples of past research, a brief overview of current research methods, and suggestions for future research. A 34-item bibliography is attached. (Author/RAA)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED197710

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

THE USERS OF INDEXES

by

Christina Beal

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Christina Beal

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

IR009096

THE USERS OF INDEXES

CHRISTINA BEAL
Research Officer, Centre for Research on User Studies, University of Sheffield

The problems of indexing and the use of indexes lie at the heart of information use in the broadest sense. They determine maximum possible use of information systems and sources (including libraries), they affect communication with users and they influence the efficient use of documents once retrieved. Because of this they should form a more important field of research than they have in the past. This paper aims to outline examples of previous work on use and users of indexes and to expand these into what seem significant areas for further study.

It begins by taking a look at past work on evaluating the effectiveness of indexing systems from the user's point of view. This is followed by sections on index searching, including person-computer interactions in on-line information systems. Past attempts at categorising index users provide some clues to both user behaviour and user needs, but the very nature of indexing systems tends to solidify the information worker's view of the world and systems are sometimes accepted as unproblematic for the user. This is often not so. Some work in the public library and information field has tackled the user problem by trying to structure index systems from users' own formulations of enquiries. This is not always practical or possible, and there is strong feeling that more instruction is needed in schools to teach children how to make full use of existing indexes and index systems, to make better use of index search skills in everyday life and to view indexes as diverse ways of structuring knowledge and information so that in future they can adopt a more flexible approach to indexes in general. A brief overview of the research methods being used in some of these current projects is followed by a few suggestions on what research on users in the 1980s and beyond might look like.

Evaluating the effectiveness of indexing systems for users

Much evaluation of the effectiveness of indexing systems for users tends to look less at the actual process of searching in, or use of, a system, and the eventual outcome of the search for the user, than at its end-product - the document or documents retrieved. Often a quantitative measure of effectiveness is taken such as the average number of items retrieved from the system per day, the daily count of successful searches where one or more items were retrieved for a user, the effort and cost of using the system or the response time. Some of these are of less concern to the user than they are to the managers of the system. Other quantitative measures such as recall and precision ratios have also been devised to take account of the number of relevant documents retrieved to satisfy the user's information needs. In the end these measures take minimal account of the user's requirements; qualitative evaluation seems essential, but as Lancaster and McCutcheon point out in discussing library services:

> In evaluation, the qualitative considerations are the most difficult ones with which to deal. In applying qualitative criteria to library services, there are two basic situations. One can be expressed in terms of "It is the service's success or failure." This is the easier one, as represented by the known item search. The user comes into the library looking for a particular bibliographic item, for which he or she has the author or title. Evaluation of this service involves a binary situation because the library either is successful or not, that is, the library either has the item the user wants at the time it is needed or it does not. The second situation, involving other kinds of services that libraries provide - such as subject searching - is not susceptible to this binary type of evaluation. In evaluating a subject search, we are faced with degrees of success which is a more difficult evaluation situation. For the user coming to a library or an information center for materials dealing with a particular subject, it's not a yes or no decision of whether some are found or not but rather the question is: how much information is found? Subject search evaluation gets into questions of relevance, of which is the most relevant or valuable material, making this evaluation... situation much more complex than the known item search.[1]

Research done at Cranfield in the 60's on efficiency of indexing systems aimed at developing an experimental method for the assessment of document relevance which took greater account of user satisfaction by evaluating the relevance to the questioner of a set of retrieved documents in response to his or her original research questions. The researchers still had to admit that their method:

> ... includes nothing which overcomes the basic problems of the meaning and determination of relevance. No one is more aware that relevance is a shifting notion, certainly between individuals and often for the same individuals at different times... We would strongly argue against [the idea that, since such measurement must be based on relevance decisions, any attempt to measure system performance is useless] for it is the very situation which an information retrieval system has to face. Users do ask questions and then accept or reject the search output in what might seem an arbitrary manner.[2]

The important question, of course, is whether or not their acceptance or rejection is truly as arbitrary as it may sometimes seem. Alternatively, even irrelevant items can be of use in focussing, or even expanding, thinking in some way.

The arguments still continue in the April 1980 issue of The Indexer, where John Regazzi takes an overview of models for the measurement of the quality of output of indexing and retrieval systems.[3] He concludes the necessity to distinguish between output and outcome in any evaluation of a system, since the goods and/or services, or outputs, of the system are used by individuals, groups or other systems for a variety of reasons and it is these uses which "ultimately determine the 'worth', 'usefulness' or 'satisfaction to the user' of the system. This seems to point strongly to the need for studies of effectiveness which more fully take account of what users think about retrieved materials and even more importantly, how they eventually put them to use. These might better demonstrate the process by which a user's information need gets translated into a search strategy, and then how the information which could be retrieved in this way is accepted or rejected as useful in fulfilling the original requirement (as perceived by the user).

Some of our work on researchers in the humanities at the Centre for Research on User Studies has given some good examples: how, for instance, key-word indexes can throw up inappropriate materials in fields where gimmicky titles aid publication prospects, in cases where

early research on a subject and a still hazy conceptualisation of the research topic make choice of index search terms extremely difficult, in cross-disciplinary research where the researcher's awareness of appropriate indexing language may be effective in his primary discipline but relatively undeveloped in the other, or where abstracts indicate a document on a desired subject but the document itself treats that subject from an inappropriate point of view.* We feel these are particularly important problems in days when words like self-renewing library and remote storage are popular; and a proposed Centre project on inter-library loans will probably also have to come to grips with questions of indexing effectiveness.

In the first six months of the Centre's existence, staff worked on producing a comprehensive file of past user studies documentation. The task proved almost impossible since the number of possible headings which could be searched in a source such as LISA were virtually endless: not only was the term itself a relatively recent one in information studies, but our own definition of the term was far broader than the one often accepted in the field. Frequently documents which were chased up because they seemed to promise a user perspective were quickly rejected when they turned out to make perhaps only passing mention of the existence of users. Our extremely slow progress on the project brought on acute depression, since we at first put it down to our own inefficiency: with time we realised the problem was one of trying to match our subject - or even more than that, our view of the subject - to the words which appeared in most indexed sources. Eventually the idea of a truly comprehensive bibliography had to be dropped altogether.

Searching Indexing Systems

Lancaster proposes a useful model for looking at actual steps in the information retrieval process, and at the factors affecting the success of each, which seems to take account not only of the user's information needs but also of the process of trying to express these needs and to have them correctly interpreted within the system.[4]

*I am indebted to my colleague, Sue Stone, for these points.

He views the process as one of linking together the indexer's conceptualisation of a document and the user's conceptualisation of the information needed; the index language "exists primarily to bring the vocabulary of the indexer and the vocabulary of the searcher into coincidence." (p.2) The factors he sees as affecting success not only include the coverage and appropriateness of the indexing system; indexing policy, quality and accuracy; and the ability of the vocabulary to describe concepts occurring in documents, but also such factors as the user's interpretation of the system's capabilities and limitations, the user's ability to describe his needs, the mode of interaction with the system, the system's guidance in formulation of a request (for example, request forms or interviews), plus the analyst's interpretation of user requirements, ability of the vocabulary to describe concepts occurring in documents and the quality of the search strategy. Obviously, the difficulties of taking account of all these factors in any one study are enormous, but it puts use of indexes more firmly within the context of the user's frame of reference and less within that of the information specialist's.[5]

One recent letter in The Guardian "Futures" feature nicely sums up the problem. It reads:

> Sir, - In his excellent article on CAFS [Content Addressable File Store] (Futures, April 30), Peter Large may have given a slightly misleading impression on the current state of techniques for retrieving technical information. The article seemed to imply that it is not currently possible to obtain technical information on solar energy by rapid, computer methods. This is, of course, not so: INSPEC, a division of the Institution of Electrical Engineers, provides this type of information on all areas of electrical and electronic technology as well as the related fields of physics and computers. Solar energy is particularly well covered and it is possible to obtain details of reports written by a specific scientist very easily.
>
> The INSPEC service, in common with other services covering other fields is accessible online via the public telephone network. - Yours sincerely.

and in reply, Peter Large, author of the original article, writes:

> The Guardian has given considerable coverage to information banks like INSPEC. Many of them, however, are librarian-orientated - you have to understand the indexing methods on which they are based. The advantage of a system of the CAFS type is that it is easily used by anyone.[6]

Wo/Man-Computer Interaction

Certainly the significance of search strategies will become infinitely more important with the increase in availability and use of online information retrieval systems and the cost of computer time. Research at Loughborough University Library involved observation of sessions where "real users" performed online subject searches using the Lockheed DIALOG system with the help of an intermediary (who actually operated the terminal). One of the main findings was that:

> The majority of customers were unaccustomed to thinking of their information needs in the terms used by the indexes or of considering the combinations which best matched their needs. They were also ill equipped to judge whether a certain number of 'hits' was good or bad and what it would cost to generate them by the alternative methods available. If the time pressure arguments are valid the only time at which they would be able to consider these issues in a relatively unconstrained way would be before the session began. Once it was underway they could well be asked to make split-second judgements on issues they only vaguely appreciate. It is normally the case that people make less than optimal decisions under these circumstances and a more considered judgement, may, for example, have led to a better search combination, or less generation of irrelevant references. It should be remembered that, in this experimental study, customers were not being charged directly for the cost of the search. They may therefore have been less conscious of the costing issue than the intermediary and therefore less prepared for rapid decisions. On the other hand, if they were to be charged directly, they may well have been even more anxious to reduce wasted time and have made even worse decisions.[7]

Main conclusions from the study were that if the customer is to achieve "the best match possible between information needs (expressed in his own terms) and what the system can provide, he must be given time to identify how this may be done." The study felt that factors relating to this process included the fact that: the problem of selecting terms and combinations is a skilled job and not simply a matter of knowing the topic in question, so that the skills of the intermediary are probably more important in making decisions in the course of a search than they are in the mere operation of the system.

These skills "develop with experience of the system and hence the combination of the customer's knowledge of his needs and the intermediary's guidance in the use of the system is essential to its successful use... The characteristics of the man-man relationship should take precedence over the characteristics of the man-computer relationship." (p.102).

The intermediary - the information specialist, librarian, indexer - should become more, not less, important in the future.

Categorising users of indexes

The most common way of categorising index users for purposes of either describing current index use or for designing new indexing systems is by dividing them up into scientists and technologists, workers in the humanities, social scientists and members of the general public or laymen. Langridge has given a good overview of distinctions between some of these user groups.[8] Within the sciences and technology he distinguishes between three broad groups - practising scientists and technologists, students and laymen - each of which has quite different indexing needs according to their basic approach to, and requirements of, the discipline. For example, "the literature used by [practising scientists and students] is mainly incomprehensible to [laymen], their needs being met by the most elementary textbooks, popularisations and discussions of scientific findings. The indexing needs of this group must obviously be considered quite independently of the other two, particularly in view of terminology problems." (p.117) In contrast, in the humanities, although there are three comparable groupings (which Langridge terms original creators, students and the 'general reader'), these are far less homogeneous and distinct than in the sciences. As examples, he points out that: the requirements of scholars in the humanities are likely to be much wider ranging than in the scie[illegible] (for instance, there is no definable limit to the range of information that a novelist may require), the interests of scholars overlap considerably with those of students, students in turn are not sharply distinguished from the 'general reader' and the term 'layman' is not really appropriate since the majority are not excluded from the study of humanities literature in the way they are from that of scientific. He concludes:

> In the sciences, then, we have a comparatively homogeneous body of knowledge, in contrast to a comparatively marked distinction between groups of users. In the humanities we have a heterogeneous set of disciplines, but no clear distinction between groups of users. Generally speaking, the task of indexing

> services in the humanities is to provide a comprehensive service for all users, rather than a series of services for different groups. There is no parallel in the humanities to the time problem in science... [and] the use of very detailed indexing in the humanities is not largely the prerogative of one group: it is not only theologians that will use a concordance to the Bible, or literary scholars a concordance to the works of Shakespeare.
>
> It was not the intention of the preceding discussion to establish the precise nature of various user groups. Only the broadest distinction between the sciences and the humanities is being considered. This is enough to show that any tendency to generalise in matters of classification and indexing from science and technology is condemned by the requirements of users, as it is by the nature of knowledge itself. (p.118)

Eugene Garfield also discusses the basic differences in information retrieval in the sciences and the humanities:[9] scientists: are 'reductionist'; try to locate theories and methods in the literature that in any way affect their own current research; are concerned with priority; want to assure themselves that no one else has already published the research they are working on; build on their own and others previous work and so have a tradition of referencing earlier work that has influenced them. The humanities are concerned with human achievements making up a cultural heritage, rather than with the physical universe; works are not superseded but cumulative; the arts and humanities are not static and an inheritance may be re-evaluated to become popular again so that one of the significant characteristics of the humanities is shifts of interest. Scholarly work tends not to develop out of previous work but as a reaction against previous perspectives, making scholars less compulsive about the literature than scientists; there is a preeminence of books over journals as sources and these sources do not become obsolete since great scholarship and criticism endure; scholars wish to retrieve actual works of art themselves; the citation tradition is not as strong as it is in the sciences.

In the early stages of the INFROSS project a list was drawn up of the main characteristics of the social sciences, relevant to information and its communication, which distinguished them from the sciences.[10] The subject matter of the social sciences (human beings interacting or acting in groups) was seen as being much less stable than that of the sciences; there was a great variety of approaches to its study; each discipline was much less clearly defined; the subject scatter of potentially relevant information was much greater; the solid factual content of most of the social sciences was quantitatively smaller and, almost certainly, less vitally important; concepts and ideas - abstract

thinking - played a larger part. This unstable subject content and strong conceptual element in the social sciences tended to make indentification of required information more difficult and: "Instead of being an easy preliminary to the searching process, identification is often an integral part of it - the information only being recognised as required when it is seen." This was seen to make browsing more important and to require information services to be more positive, aiming at exposure to information of potential relevance. Finally, discovery in the social sciences was seen as being a different phenomenon from in the sciences and the need to establish priority in an area through publication much less important; social science discoveries and ideas might become outdated, but were not superseded since they came into being in a social context and, for that reason alone, were of continuing interest to social scientists. Barbara Skelton, in a later paper, compared the results of a number of studies (including INFROSS) in order to draw up generalized profiles of scientists and social scientists as information users.[11]

How do all these group characteristics affect index search behaviour and consequently indexing needs? The research which should best offer some clues to these questions are studies into information needs, both documentary and otherwise, in particular disciplines, subjects or professional groups - INFROSS,[12] Bernal's early work on scientists,[13] CRUS' own study of humanities scholars,[14] Imperial College's 1975 study on communication and information needs of earth science engineers,[15] Grayson's overview of information users in local government,[16] OSTI's report on the use of information sources in business studies[17] - to name only a very few at random. But in these, unfortunately, problems of information search strategy - at least in terms of document searching - have not often been tackled in depth from the point of view of conceptualisation and categorisation of information problems, placing of subjects within broader subject areas, perceptions of fields of knowledge and relationships between them and finally translation of these into searching of indexes and information sources. In general, there tends to be more of such information coming out of the work being done on specific indexes or classification schemes since, rightly or wrongly, in studies of information need indexing has rarely been seen as problematic.

One of the most interesting examples of such work is a project undertaken at the Open University in the early 70's concerned with the "problems of indexing and classification of social science literature, with particular reference to the sociology of education."[18] The aim was to develop indexing techniques appropriate to the field through the cooperation of a group of subject experts, and by compiling a bibliography and index of a sample of relevant material.

The researchers painstakingly trace the course of their work through its early stages, where they were looking at the appropriateness of the PRECIS indexing system for their purposes, to their efforts at overcoming the problems they found with this system ("What it took us some time to recognise is that, irrespective of the intellectual content of modern analytic approaches to indexing and classification, there is a conceptual orientation built into the fundamental principles underlying the mechanics of schemes employing such approaches which defeats all attempts to sensitize a scheme to a range of alternative approaches." p.15), and finally through their attempts to find an alternative rationale for indexing which would be "more in line with the concerns of users searching the literature" - which would allow the user to respond in terms of his or her own conceptual orientation. In the process they explore many important ideas about meaning and conceptualisation as these relate to indexing and index use and, as the following quotes show, by so doing make us reflect more fully about information science research itself:

> ...the problem lay in assuming that the literature in a given field comprises a universe of knowledge which the information scientist may map by analysing the subject matter to which documents refer in terms of the elements they comprise, and the relationships amongst them, and by defining this as the universe of public knowledge. Whether a scheme provides for analysis of content or meaning in terms of concepts, words, word fragments or any other element, and whether the bits are put back together ab initio by the system designer, or at the search stage by an information officer or user, the same implicit assumption seems to underlie all the schemes. The assumption is the objectivist one that meanings are to be regarded as entities which remain constant when divorced from context (Vol. 1, p.23)

> ...the apparently dominant assumption in information science that the meaning of information is independent of the language in which it is expressed, and of the context of its use in terms of the perspectives of author and user... appears to run right through the process of research and development in information science. (vol. 1, p.25)

> ...the meaning of terms used to describe the content of documents does not necessarily inhere in the phenomena under investigation, nor in the documents describing them, but may depend upon assumptions external to both. Therefore the framework of an effective indexing or classification scheme must similarly represent factors external to the subject matter with which it deals. In other words, material must be presented in terms of the dimensions of knowledge by which it comes to be structured [both by the author and by the user]. However, to the extent that meanings remain partly implicit, both because they have multiple sets of associations (everyday and technical) which cannot always be enumerated, and also because a writer may assume that he can rely upon a certain background of experience amongst those he is addressing, we must assume a measure of user involvement as an integral element in system development, use and maintenance, for which information science skills will be able to devise no alternative. In other words, both organisation and retrieval processes will be situations not of the system acting for the user, but of the user as interacting with and part of the system. (vol. 1, p.27)

> We have argued that the imposition of structure confers meaning, and that meaning determination is not properly a task which a system can undertake for its users, because it cannot enter into their minds. However, if we are to facilitate the literature searching process we shall necessarily impose structure. In practice, therefore, the effective system will be one in which structure (1) is kept to a level which does not intervene seriously in the research process, and (2) is sensitive to the range of user perspectives rather than rationalising perspectives away or ignoring them. (vol. 1, p.281)

These problems are ones of which information researchers are becoming increasingly aware. At the 39th FID Congress in Edinburgh in 1978, Belkin explained how a new approach to theory and experiment has begun to emerge. It shifts away from looking at texts and operations upon texts to looking at users and interaction of users with texts.[19] He sees past work as having been "primarily constrained within the limits of existing systems of document description and retrieval" and being further limited "by ignoring or taking for granted, user-oriented issues such as relevance [so that] the effect of theory and experiment on practical information work has been close to negligible." He feels the new approach is one which could make theory and experiment more relevant to the concerns of practising information workers. More importantly, it could perhaps help practitioners to challenge accepted views held about the systems which they operate.

Some thinking on the information worker's view of the world

Hatt, in his book on the reading process,[20] writes: "It is possible that many of the difficulties encountered by users of library catalogues and bibliographic tools can be ascribed to the fact that readers are trying to find texts, through the medium of an apparatus designed to effect re-finding." (p.67) In dividing up a field of knowledge, a category or term has to be defined with reference to other categories found in the field, but in searching

for an item of knowledge, there may not be any clear view of the field in the user's mind, let alone one so structured that it throws up categories into which the sought item can be neatly slotted. He argues that the nature of language itself is that it allows us to take our acts of classification for granted; what the reader is being asked to do in using indexing and classification systems is to make conscious and explicit these organising and distinguishing activities usually taken for granted. It is as though in asking for a cheese sandwich made with brown bread, he says, the reader is made to say 'I want a sandwich which is a kind of snack, made with bread which is a cereal food, and butter which is a dairy food, and cheese which is another dairy food, and I want the bread to be brown and not white.'

The danger is that in time the information specialist's view of this categorisation of information, represented in both indexing and classification schemes, comes to seem normal and natural, instead of one which is problematic for users until they learn how to negotiate it and frame their questions in the language of the system.[21]

To date, work which seems to come closest to dealing with this problem has been, quite appropriately perhaps, in the public library and public information fields. A recent Danish project,[22] for instance, has tried to solve the problem of fiction arrangement and classification in public libraries (usually based solely on either author's name or genre-terms) by developing a scheme more closely related to readers' formulations and criteria of choice. The pilot scheme was based on an analysis of 300 user/librarian conversations recorded under everyday library conditions, and the users' categorisations of fiction subsequently characterized according to four dimensions: subject-matter; frame or setting in time or place; author's intention (usually emotional experience or cognition/information) and accessibility both in terms of readability and physical characteristics: these are further subdivided into classes. A reader's enquiry could be slotted into this scheme at any point. The scheme is not exclusive since all characteristics may be found in the same book; a book can be classed in several places in the system; since there is no logical order in the established classes there is no hierarchy of classes and the librarian can enter the system anywhere without first having to place the subject in some superordinate class and then search on for a subordinate class. For example, a reader could ask in terms of accessibility - for an easy book, in terms of emotional experience - for an exciting book, in terms of frame - for a war novel, in terms of cognition/information - for a polemical book, and with any

or all of these requirements, be easily directed to suitable reading material. This gave the scheme a great amount of flexibility, as well as an effective means for librarians to structure interviews with their readers when trying to answer enquiries. A similar scheme applied to music score libraries, for example, would help the piano-playing user who wants something by Brahms which is relatively easy to play: the traditional arrangement of scores by composer, subdivided by music form, can prove less than helpful.

Community information research too has had to try and deal with how people talk about or characterize their everyday problems and information needs: this becomes vital not only for the researcher trying to find out more about needs,[23] but also for the advice or information agency aiming to translate these into user-appropriate filing systems, resource centres and even statistical records of enquiries made to the service.[24]

In talking about book indexes, Edward de Bono has brought up some of the same issues:

> As for indexing: do we feel that we are in the communication or in the thinking business? Do we work to follow the thinking process of the person who has produced the book, or the consulter, or our own?... An index may be indicative or creative - do you lead readers to find what they expect to find or do indexers have a brief to look at things in a new way? We have no chance to explain to readers why we have viewed concepts as we have.[25]

Teaching index search skills in schools

Michael Marland has listed three uses or functions of document indexes which he feels strongly should be taught in schools:[26] the first is "to reveal the scope, level and detail of a book and thus its suitability for the purpose in hand"; the second, the index "unorganizes" the author's organization and allows the reader to make his own reorganization; third, the index is a recall device. He feels that teaching reading is insufficient without also teaching how to make use of the ability to read, to analyse and reorganize printed information or, as he calls it, "reading for learning" - equipping children to use books as learning sources. Part of this teaching should involve the methodical teaching of study skills; and this should include the study of classification within a book and within a chapter, as well as how to use an index (adding in passing that "it would be helpful if children's books had good indexes").[27] He includes in this the need for special instruction in the use of the library.

One thing we know little about is how far children conceive the organization of material in a library, on the shelves, in a card catalogue and elsewhere as static representations of bodies of knowledge and existing documentation of subjects. I was struck recently by a reply from a youngster in one researcher's study which clearly indicated there was little awareness that skills and practices in using one particular library - in this case the school library - could be transferred to the use of another library such as the local public library. That school library was seen as a one-off place; the danger is that that youngster's perceptions of a library and of books as representations of a world of knowledge became governed by what he saw as available in that one library.

The full implications of this become clearer in the case of another school library, arranged by Dewey classification, where, for historical reasons, most of the books on politics were located in the history, economics or general studies sections. Without proper guidance on the principles of that indexing system and on relationships between subjects, it is possible that children in that school be left unaware, and consequently "deprived", of one particularly important field of knowledge.

The importance of systems of indexing and classification as representations of certain views about the organization, and even the existence, of bodies of knowledge has to be fully realized: unless youngsters come to see these representations as just that - representations - they will be seriously hampered in acquiring a flexible approach to information and to the use of tools for accessing that information.[28]

The importance of index search skills in everyday life

The extent of the difficulties experienced by many members of the public in categorising their information problems and using index search skills in order to trace everyday information is probably underestimated. A British Library funded schools project by Terence Brake carried out with fourth-year pupils in South Hackney[29] has provided new insight into the problem on a very practical level. The project aimed at exploring the sort of social or community information which young people need both while at school, and when they leave, as well as the information skills which pupils need to acquire before leaving school in order to make adequate use of information in the management of their everyday problems. A community information unit was set up in the school library, including a local information card index, telephone directories, Yellow Pages, a literature file and so on.

One of the main teaching tasks was seen as developing perceptions of the relevance, adequacy and critical understanding of information. For example, one teaching exercise - a 'Film Night' worksheet - took a trip to a West End cinema as its theme and broke this leisure activity into its component parts. The exercise used copies of Time Out and photocopied pages from evening newspapers to supply information on choice of film; London Transport (tubes and buses), British Rail and A-Z maps to provide information on travel facilities and the location of cinemas, and the telephone to find out other information such as price of seats. All these activities required the use of basic index search skills which to many of us may seem too simple for words, but which were far more problematic for many of the pupils. One notable comment on the exercise was that "The use of maps was stimulating to most pupils - 'This is where I live, look!' - although the London bus map, in particular, was daunting to some pupils - 'How are we supposed to look at that?'" Another exercise, based around use of a booklet called "What's on in Hackney" also aimed not only at making pupils aware of the wide variety of local information sources but also at getting them to practise simple searching skills. They were asked to guess under which headings in the What's On booklet they would find various organisations and then, after guessing, they were asked to search the booklet to see if their guess matched up with the actual location of the listing of the organisation. Few of the pupils had seen the booklet before and the questionnaire apparently prompted questions such as "What's national radio?" and "What's a circular?" The researcher's message about the purposes of this headings exercise seems important to bear in mind: it was "to illustrate the subjective nature of the organisation of information, and to show how it may be necessary to explore a document beyond one's expectations because whoever organized that information did not see a particular item fitting into the scheme of things in the same way as oneself. In this context there were no right or wrong answers; however, the exercise was often interpreted in this way [by the pupils] - 'Ha, ha you got it wrong'. Others would be jealously guarding their guessed headings. This made the exercise appear to be more difficult than it actually was." (p.30)

Such a lack of skills for coping with basic indexed sources of information like street maps and directories (always bearing in mind that many of the difficulties experienced with sources like telephone directories probably stem from how badly they are indexed) casts doubts about the possibilities for many members of the public to effectively make use - in their own homes

or in public libraries or post offices - of more complicated categorisations of information such as those now being used in the Post Office's Prestel system. This problem certainly needs more attention than it seems to be getting at the moment. The work on PRESTEL by Cathy Sullivan and David Oliver at Aslib Research and Consultancy Division, some of it to do with the structuring of local files of community information and getting public reaction to them,[30] seems to be heading in the right direction.

Research methods

A very brief overview of current research would show that interviewing and questionnaire methods are beginning to be used more extensively in assessing user reactions to specific indexing systems and services. The Courtauld Institute of Art's work on users' reactions to RILA (Repertoire International de l'Art) and Ken Bakewell's work at Liverpool Polytechnic on PRECIS are two good examples of this. Out of all these surveys a clearer picture should hopefully begin to emerge about index use and index users as part of a normal, continuing and increasingly accepted, as well as increasingly important, process of service evaluation.

In looking at the problems of interview and questionnaire methods for studying library use Lancaster and McCutcheon write:

> Ideally, we would like to be able to measure user behavior in the library by unobtrusive means to counteract the danger that behavior may be altered if the individual knows s/he is being observed. In most aspects of library usage, there is no way of measuring the performance unobtrusively... User success [or failure] can only be determined by asking the individual patron. (p.18)

These points would seem to apply equally to observation of index use in general. The authors also point out that interviews can vary in level of obtrusiveness according to when they are conducted: those conducted after use or after an index search being the least obtrusive, and perhaps the most appropriate, running interviews (in which the researcher questions a user throughout a search) being the most obtrusive. Obviously, asking users to remember, conceptualise and describe past behaviour is similarly problematic. However, the authors seem to feel that the increasing mechanization of libraries could offer an ideal solution to unobtrusive measurement (or what they incorrectly call "objective measurement") and evaluation of services: "If, rather than consulting a card catalog, a user comes to a terminal, to search

an on-line catalog of a library or a group of libraries, performance can be monitored completely unobtrusively." Indeed, it may in future be possible to record and analyse both simple and complex search patterns in this way; but it is necessary to be cautious about the extent, and certainly the quality, of this information: no machine can record human thought processes or judgements...yet.

There seems to be more use made of experimental procedures in this field than in others. Recent experiments include some of the work of the Centre for Catalogue Research at Bath University, such as their Birmingham Polytechnic Central Library catalogue experiment in which a short-entry version of the normal catalogue replaced the full version at all usual catalogue positions in the library for four weeks in the middle of term, and only one copy of the full version was kept at the enquiry desk. Objectives were to find out how, where and why the short-entry catalogue failed to satisfy users' needs and when and why library staff required the full version. A related piece of research at City University looked at whether a simple finding list catalogue was easier to consult than a full catalogue. In 1978 the British Library awarded a grant to Manchester Polytechnic to set up and monitor development and use (including motivations of those using the service and their reactions to it) of an experimental on-line bibliographic information service in Lancashire County Library, and recently another award to the Courtauld Institute to establish a RILA experimental on-line service and to evaluate its usefulness through detailed records of individual searches, and so on.

Much more work, however, still needs to be done on how, for example, the indexing needs of library users could be studied or how the needs of book index users might be assessed.

Future research on users

Brookes, in a paper given at the 1978 FID Congress in Edinburgh,[31] presented a view of present document-retrieval systems (he calls them Stage I systems) as "mechanized replicas of manual systems" concerned with the organization and mechanization of documentation. Stage II systems, he envisaged, would present users directly with synthesized information of 'knowledge structures'; they would be concerned with the organization and mechanization of information. He views the present Stage I systems as necessary and ideal testing grounds for empirically studying "interactions

between new theory, new technological developments and users". But, he warns:

> The intimate cognitive processes by which humans synthesize the information they absorb have defied analysis since the time of Plato and Aristotle. Yet the need to make coherent sense of the world's increasingly rapid generation and cumulation of data-banks and data-bases confronts all present systems with this same ancient problem. The present rate of generation of information far exceeds our ability to synthesize it all for use.

Studies of users must adopt a new theoretical basis for research, or a new research principle. "This principle states that, though the phenomena of the physical world can be adequately described in terms of the fundamental categories of space/time and matter/energy, for the adequate description of cognitive phenomena we need to admit also the additional fundamental categories of information/knowledge..."

At the simplest level there are countless questions to be asked, and answered, about the use of mechanized information systems - not least the way that people can cope with the lack of a physical representation of large amounts of information - in the form of complete, physically-indexed, documents, sets of documents, shelves or files of handleable material. Could a paperless library - even a paperless office filing system - really work? How important is it for us to handle, manipulate, reorganize information in some physical format - on pieces of paper, in piles on our desk, in files in a drawer or on a shelf? Some may be able to retain a mental picture of a body of information; others may need a concrete representation of it. What is the nature of any link between ability to absorb information and the possibility of accessing, organizing and re-organizing it manually? How indeed can one rationally conceive of indexing and accessing such immense stores of information as are becoming available, with current restricted knowledge of human thought processes. Has our technology really outstripped our ability to make efficient and effective use of it? Do people want to have access to the world's total store of recorded information on any subject: if they do, can they cope with it once they get it? At times, systems like EURONET,[32] at least to an amateur, can seem infinitely frightening.

The difficulties for both indexers and users posed by the increasingly international nature of these systems are enormous, and as Noël wrote in another paper for the 1978 FID Congress,[33] "... the prospects of fruitful cooperation between linguists and documentalists are much brighter than ever before... In information science, the problem of the masses of information to be processed speedily is one with which linguists can hardly be concerned. The area in which linguists and documentalists can cooperate is that of users'

language needs: cross-disciplinary problems of syntax, of terminology, and of the 'language-barrier'."

The British Library has recently funded UMIST to study the feasibility of a terminological data bank,[34] or term bank, in the UK. (The project description defines this as "... a body of specialized terms, with definitions, stored in a computer; it can be used as a dictionary for direct consultation, as a control instrument to ensure consistency of terminology and as a basis for the production of printed dictionaries.") As part of the project, potential users in government departments, industry, library and documentation centres, publishers and translators' organizations will be asked to comment on a number of alternative models for such a service. It seems that not only tomorrow's technology, but also techniques for information specialists to cope with it, may just about be here, but studies of users haven't yet caught up with yesterday.

REFERENCES

1 Lancaster, F.W. and McCutcheon, D. 'Some achievements and limitations of quantitative procedures applied to the evaluation of library services' in Ching-chih Chen, ed., Quantitative measurement and dynamic library service. London, Oryx Press, 1978. p.16.

2. Cleverdon, C. et al. Factors determining the performance of indexing systems. Vol. 1. Design. Cranfield, 1966. p.16.

3. Regazzi, J.J. 'Evaluating indexing systems: a review after Cranfield'. The Indexer 12 (1) April 1980. 14-21. Regazzi also quotes a major criticism of these models voiced by Soergel - that the worth of articles retrieved gives no clue to the possible value of those not retrieved but within the system (Soergel, D. 'Is user satisfaction a hobgoblin?' JASIS 27 (4) 1976. 256-9)

4. Lancaster, F.W. Vocabulary control for information retrieval. Washington, Information Resources Press, 1972.

5. There seems to have been more work on search strategies in catalogue use in the past than in other areas. For example, in 1972 Lipetz reported an extensive study, involving more than 2,000 interviews and lasting one year, of use of the main catalogue at Yale University Library. This covered initial search objectives, starting clues, intended approaches and successes and failures. Lipetz, B.A. 'Catalog use in a large research library'. Library Quarterly 42 January 1972. 129-139.

6. 'Futures: the world of science and technology'. The Guardian May 8 1980. p.24.

7. Baker, C.A. An investigation of man-computer interaction in an online bibliographic information retrieval system. Loughborough University of Technology Library, May 1977. p.100.

8. Langridge, D.W. Classification and indexing in the humanities. London, Butterworths, 1976. 117-8.

9. Garfield, E. 'Is information retrieval in the arts and humanities inherently different from that in science: the effect that the Institute for Scientific Information's Arts and Humanities Citation Index is expected to have on future scholarship.' LQ 50 (1) January 1980. 40-57.

10. Bath University of Technology Library. Investigation into information requirements of the social sciences. Report on the preliminary stage. April 1968. 11-13.

11. Skelton, B. 'Scientists and social scientists as information users: a comparison of results of science user studies with the investigation into information requirements of the social sciences.' J. Lib. 5 (2) April 1973.

12. Bath University Library. Investigation into information requirements of the social sciences. 1971.

13. Bernal, J.D. 'Preliminary analysis of pilot questionnaire on use of scientific literature'. Royal Society. Report of the Royal Society information conference. London, 1948. 101-2, 589-637.

14. Corkill, C. and Mann, M. Information needs in the humanities: two postal surveys. University of Sheffield, Centre for Research on User Studies, 1978. (CRUS Occasional Paper 2)

15. Gralewska-Vickery, A. and Roscoe, H. Earth science engineers communication and information needs. Final report. Imperial College Rock Mechanics, 1975. (OSTI Report No.20)

16. Grayson, L. Library and information services for local government in Great Britain. London, L.A., 1978.

17. Dews, J.D. The use of information sources by teachers and research workers in the field of business studies. Report to the Office for Scientific and Technical Information. Manchester Business School, 1970.

18. Swift, D.F. et al. A case study in indexing and classification in the sociology of education: development of ideas concerning the organisation of material for literature searching. Report for the period September 1970 to June 1973. Milton Keynes, The Open University, 1973.

19. Belkin, N.J. 'The cognitive viewpoint: a new basis for research in information science.' 39th FID Congress. University of Edinburgh, 25-28 September 1978.

20. Hatt, F. The reading process: a framework for analysis and description. London, Clive Bingley, 1976.

21. ... internalisation of the orientation assumed by current indexing systems may be expected to reduce system designers' and operators' capacity to 'see' a document in any other way. This has a bearing also on the appropriateness of the criteria of evaluation which are employed in investigating the effectiveness of their systems, and which users become conditioned to regard as valid. In other words, it seems probable that the situation operates to legitimise current practices." (Swift et al. A case study in indexing and classification in the sociology of education. The Open University. 1973. Vol.1, p.25.) The extreme difficulty of balancing up the advantages and disadvantages of different types of indexing schemes to users' requirements is clearly outlined in chapter 8 of Langridge's Classification and indexing in the humanities (London, Butterworths, 1976).

22. Pejtersen, A.M. 'Design of a classification scheme for fiction based on an analysis of actual user-librarian communication, and use of the scheme for control of librarians' search strategies.' Paper to be presented at the 2nd International Research Forum on Information Science (SIRE), Copenhagen, August 1977.

23. The Baltimore study is a major example of such research: Warner, E.S. et al. Information needs of urban residents. U.S. Department of Health, Education and Welfare, Office of Education, Bureau of Libraries and Learning Resources, 1973.

24. The National Association of Citizens' Advice Bureaux has recently been awarded a grant by the British Library for a preliminary investigation of the organization and classification of community information.

25. Edward de Bono, reported in The Indexer. October 1978. p.62.

26. Marland, M. 'Reading to learn and using the index'. The Indexer 11 (2) October 1978. 68-9.

27. Marland, M. 'The role of the information book in the school'. The Indexer 11 (1) April 1978. 45-6. In her review of indexes of children's books in Australia in the April 1980 issue of The Indexer , Brenda Miller writes: "I cannot stress too strongly that for many children using an index is a laborious task, and the index should therefore be made both as simple, and yet as comprehensive as possible. It is daunting, even for a literate adult, to be faced with a large number of page references under a single heading, and many children will give up at this stage. If they cannot find the information they need because the index is inaccurate, or because vital topics do not appear in it, they will give up using indexes. As I wrote in 1973, 'They have a right to expect that the indexes they consult should be accurate, comprehensive and simple.' Unfortunately I can find no evidence that this state of affairs is any closer now than it was then." 'Indexes of children's books in Australia: a second look'. The Indexer 12 (1) April 1980. 29-33.

28. A recent point made in 'Comment' (LAR 82 (3) March 1980) about the need for a new general classification: "Multi-faceting is the great problem, increasing multifold with the growth of inter- and multi-disciplinary books. In the end, perhaps, it is the user who has to become alert to the unavoidable limitations of any classification system and allow the mind as well as the eye to browse."

29. Brake, T. The need to know: teaching the importance of information. Final report for the period January 1978 to March 1979. British Library Research & Development Report No. 5511. January 1980.

30. Reported in Library and Information Research News. 2 (10) June 1980. p.14.

31. Brookes, B.C. 'Informatics as the fundamental social science. 39th FID Congress, University of Edinburgh, September 1978.

32 A European Community network presently being developed to provide access to scientific, technical, social and economic information.

33 Noel, J. 'Syntax, semantics, and pragmatics in the automatic analysis of texts.' 1978 FID Congress, Edinburgh, September 1978.

34. Project Head: Professor J.C. Sager, Department of European Studies and Modern Languages, UMIST.